Recent Results on Non–Adiabatic Transitions in Quantum Mechanics

George A. Hagedorn and Alain Joye

Abstract. We review mathematical results concerning exponentially small corrections to adiabatic approximations and Born–Oppenheimer approximations.

Introduction

The goal of this paper is to review recent results on exponentially small non-adiabatic transitions in quantum mechanics. In Section 1, we provide background information about adiabatic approximations. In Section 2, we discuss the determination of non-adiabatic scattering transition amplitudes. In Section 3, we describe the time development of exponentially small non-adiabatic transitions. In Section 4, we turn to exponentially accurate Born–Oppenheimer approximations. Then in Section 5, we discuss the determination of non-adiabatic corrections to Born–Oppenheimer approximations in a scattering situation.

1. Adiabatic Background

The adiabatic approximation in quantum mechanics concerns the time-dependent Schrödinger equation when the Hamiltonian depends on time, but varies on a very long time scale. Mathematically, this situation corresponds to the singularly perturbed initial value problem

\[ i \epsilon \partial_t \psi_\epsilon(t) = H(t) \psi_\epsilon(t), \quad \psi_\epsilon(0) = \phi(0), \]

where \( t \in \mathbb{R} \), \( \epsilon \) is a small parameter, and \( \psi_\epsilon(t) \) belongs to a separable Hilbert space \( \mathcal{H} \).

In a simple situation, the adiabatic approximation relies on two basic assumptions. The first is a regularity condition on the Hamiltonian:

\[ \mathbf{R}: \text{The Hamiltonian } H(t) \text{ is a bounded self-adjoint operator on } \mathcal{H} \text{ that depends smoothly on } t. \]

1991 Mathematics Subject Classification. Primary 81Q05, 81Q15; Secondary 81Q55, 81Q20.

Key words and phrases. Molecular quantum mechanics, adiabatic approximations, Born–Oppenheimer approximations.

Supported in part by NSF Grant DMS–0303586.
We denote the unitary propagator corresponding to (1.1) by \( U_\epsilon(t, s) \). It satisfies
\[
(1.2) \quad i \epsilon \partial_t U_\epsilon(t, s) = H(t) U_\epsilon(t, s), \quad U_\epsilon(s, s) = I
\]
for all \((t, s) \in \mathbb{R}^2\).

The second assumption is that there is a gap in the spectrum of \( H(t) \) at all times \( t \):

**G:** The spectrum of \( H(t) \), denoted by \( \sigma(t) \subset \mathbb{R} \), can be decomposed as
\[
\sigma(t) = \sigma_1(t) \cup \sigma_2(t), \quad \text{with} \quad \inf_{t \in \mathbb{R}} \text{dist}(\sigma_1(t), \sigma_2(t)) = g > 0.
\]

Under this hypothesis there exists a self-adjoint spectral projection \( P(t) \) corresponding to \( \sigma_1(t) \), which depends smoothly on \( t \).

In this situation, the Adiabatic Theorem of Quantum Mechanics says the following:

**Theorem 1.1.** Let \( H(t) \) satisfy **R** and **G**, and let \( P(t) \) be as above. Then, for small \( \epsilon \),
\[
(1.3) \quad \| (\mathbb{I} - P(t)) U_\epsilon(t, s) P(s) \| = O(\epsilon),
\]
for any \((t, s)\), where the error term is sharp.

In other words, for small \( \epsilon \), the quantum evolution follows the isolated spectral subspaces of the Hamiltonian, up to an error of order \( \epsilon \). We call the error term the transition amplitude between the subspaces \( P(s)\mathcal{H} \) and \((\mathbb{I} - P(t))\mathcal{H} \).

One modern proof actually yields a stronger result. Let us sketch it for later reference. Let \( V(t, s) \) be the \((\epsilon\text{-dependent})\) unitary evolution defined by
\[
(1.4) \quad i \epsilon \partial_t V(t, s) = (H(t) - i\epsilon[\partial_t P(t), P(t)]) V(t, s) \quad \text{and} \quad V(s, s) = I.
\]

It is not difficult to verify the intertwining property
\[
(1.5) \quad V(t, s) P(s) = P(t) V(t, s), \quad \text{for all } (t, s) \in \mathbb{R}^2.
\]

A generalized integration by parts procedure then establishes the estimate \( \| U_\epsilon(t, s) - V(t, s) \| = O(\epsilon|t - s|) \), which implies the result.

Historically, the Adiabatic Theorem goes back to Born and Fock [9] who proved it in 1928 for matrix valued Hamiltonians with non-degenerate eigenvalues. In the fifties, Kato [31] proved the statement for general Hamiltonians with an isolated eigenvalue \( E(t) \). Thirty years later, Nenciu showed the isolated eigenvalue \( E(t) \) could be replaced by an isolated piece of spectrum \( \sigma_1(t) \). The extension to unbounded operators was proven by Avron, Seiler and Yaffe in [2].

Although we shall always assume the gap hypothesis **G**, there are recent versions of the adiabatic theorem that hold without it. Instead, these results require the existence of a sufficiently regular, finite rank family of projectors \( P(t) \) that satisfy \( P(t) H(t) = E(t) P(t) \) for all \( t \). Here \( E(t) \) is an eigenvalue of \( H(t) \), and \( P(t) \) is a spectral projector of \( H(t) \) for almost all \( t \). See e.g., [10], [1], [43]. In this situation, the transition amplitude can go to zero arbitrarily slowly as \( \epsilon \) tends to zero.

Since the early thirties, it has been known that the adiabatic theorem could be substantially improved under certain circumstances. Zener [45] considered (1.1) for the \( 2 \times 2 \) matrix Hamiltonian
\[
(1.6) \quad H(t) = \frac{1}{2} \begin{pmatrix} t & \delta \\ \delta & -t \end{pmatrix},
\]
whose spectrum is $\sigma(t) = \{ \pm \frac{1}{2} \sqrt{t^2 + \delta^2} \}$. For this Hamiltonian, (1.1) can be solved exactly in terms of parabolic cylinder functions, for any $\delta$ and any $\epsilon > 0$. In particular, in the scattering regime where the initial and final times $s$ and $t$ tend to $-\infty$ and $+\infty$ respectively, the transition amplitude takes the simple form

$$A(\epsilon) := \lim_{s \to -\infty} \lim_{t \to +\infty} \| (I - P(t)) U_\epsilon(t, s) P(s) \| = e^{-\pi \delta^2/(4 \epsilon)}.$$  

Zener’s analysis shows that the transition amplitude decreases from its order $\epsilon$ value for finite times $(s, t)$ to an exponentially small value in $1/\epsilon$ in the scattering limit. Shortly after Zener’s result, Landau [32] argued that (1.7) should also hold for more general analytic $2 \times 2$ real symmetric Hamiltonians whose non-degenerate eigenvalues $E_1(t)$ and $E_2(t)$ displayed an avoided crossing with gap $\delta$, i.e.,

$$E_2(t) - E_1(t) \simeq \sqrt{(t - t_0)^2 + \delta^2}.$$  

Here $\delta$ is a small, but fixed, extra parameter in the problem. These considerations gave rise to the famous Landau–Zener formula, which showed that non-adiabatic transitions between spectrally isolated subspaces belonged to the realm of exponential asymptotics.

This Landau–Zener mechanism is the main concern of this review.

2. Exponentially Small Transitions in the Adiabatic Approximation

From a mathematical point of view, the first exponential bounds on transition amplitudes in a general framework were obtained in [27]. If we assume for simplicity that the Hamiltonian is bounded, this result requires the following regularity and scattering hypotheses:

A: There exists $\mu > 0$ such that the map $t \mapsto H(t) \in \mathcal{L}(\mathcal{H})$ has an analytic extension to the strip $S_\mu = \{ z \in \mathbb{C} : |\text{Im}(z)| \leq \mu \}$.

S: There exists $\nu > 1$, two bounded self-adjoint operators $H(\pm \infty)$, and a constant $c$, such that as $t \to \pm \infty$,

$$\sup_{|s| \leq \mu} \| H(t + is) - H(\pm \infty) \| \leq \frac{c}{\langle t \rangle^\nu},$$

where $\langle t \rangle = \sqrt{1 + t^2}$.

**Theorem 2.1.** [27] Assume $H(t)$ satisfies A, S and G. Then, there exist $C$ and $\Gamma > 0$ such that $A(\epsilon) \leq C e^{-\Gamma/\epsilon}$, for small $\epsilon > 0$.

The original proof in [27] establishes properties of solutions to the Schrödinger equation for complex $t$, by using complex WKB techniques. Similar results were subsequently proven using other methods. See e.g., [38], [41], [28], and [35].

A rigorous asymptotic formula for the transition amplitude across the gap of a two-level system was proven in [25]. In addition to the assumptions above, the proof required supplementary hypotheses on the behavior in the complex plane of the so-called Stokes lines of the problem.

This hypothesis, that we call SL, is complicated, and we shall not attempt to describe it here. It is typical of the complex WKB method and is discussed in detail in [25].
Theorem 2.2. Let
\[ H(t) = \frac{1}{2} \begin{pmatrix} Z(t) & X(t) + iY(t) \\ X(t) - iY(t) & -Z(t) \end{pmatrix}, \]
be a 2 × 2 hermitian matrix that satisfies assumptions \( A, S \) and \( G \). Further assume that the analytic extension of \( \rho = \sqrt{X^2(t) + Y^2(t) + Z^2(t)} \) has only one conjugate pair of zeros, \( \{ z_0, \bar{z}_0 \} \), in \( S_{i\epsilon} \), and that the supplementary hypothesis \( SL \) is satisfied. Then, there exist \( G \neq 0 \) and \( \gamma > 0 \), such that, as \( \epsilon \to 0 \),
\[ A(\epsilon) = G e^{-\gamma/\epsilon} (1 + O(\epsilon)), \]
where \( \gamma = \left| \text{Im} \int \rho(z) \, dz \right| /2 \), and \( \zeta \) is a loop based at the origin which encircles \( z_0 \).

Remarks 2.3. 1) The point \( z_0 \) is a complex crossing point for the analytic continuations of the eigenvalues of \( H(t) \), and the decay rate \( \gamma \) is the one predicted by Landau.
2) The prefactor \( G \) is equal to one for generic real-valued symmetric Hamiltonians. It can take other values for generic hermitian matrices, as was independently discovered by Berry [3].
3) Several variants and improvements of this theorem have been proven. See [22], [28], and [23]. Extensions to more general scattering systems were proved in [29], [34], and [24]. In particular, [23, 24] show that in an avoided crossing regime, hypothesis \( SL \) is automatically satisfied. See also [30].

These results were first proven by means of the complex WKB technique. Another popular and fruitful method in exponential asymptotics is the so-called optimal truncation method. In this method, one first derives an asymptotic expansion of the desired quantity in powers of \( \epsilon \). Next, one carefully estimates the difference between the exact quantity and the truncation of the expansion after \( n \) terms. This error estimate depends on both \( n \) and \( \epsilon \), and if it is bounded by \( c_0 \epsilon^n n! \epsilon^n \), one can choose \( n \) to depend on \( \epsilon \) in an optimal way to obtain an error estimate that is \( O(\epsilon^{−\gamma/(\epsilon^g)}) \). The optimal \( n \) satisfies \( n \sim g/\epsilon \), where \( g > 0 \) is sufficiently small.

Nenciu [38] first implemented these ideas in a general adiabatic context. His work followed contributions by Garrido [12] and Sancho [40] in the sixties. The optimal truncation technique was also used in [28], and an elementary exposition of it is provided in [19]. More recently, [39] and [44] further adapted and generalized these ideas to deal with space-time adiabatic theory.

Let us briefly present this circle of ideas through the construction used in [28]. For a Hamiltonian \( H(t) \) satisfying \( R \) and \( G \), we set \( H_0(t) = H(t) \), \( P_0(t) = P(t) \) and \( K_0(t) = [\partial_t P_0(t), P_0(t)] \). Then, we iteratively construct a sequence of self-adjoint operators as follows. For any integer \( q > 0 \) and \( \epsilon \) sufficiently small, we define
\[ H_q(t, \epsilon) = H(t) - i \epsilon K_{q-1}(t, \epsilon), \quad \text{and} \]
\[ K_q(t, \epsilon) = [\partial_t P_q(t, \epsilon), P_q(t, \epsilon)], \]
where \( P_q(t, \epsilon) \) is the spectral projector of \( H_q(t, \epsilon) \) corresponding to \( P_0 \) in the limit \( \epsilon \to 0 \). Perturbation techniques ensure that this scheme is well-defined for any \( q \), provided \( \epsilon \) is small. Furthermore, hypothesis \( S \) implies
\[ \lim_{t \to \pm \infty} P_q(t, \epsilon) = P(\pm \infty), \quad \text{for any fixed } \epsilon. \]
We next introduce the evolution \( V_q(t, s) \) associated with \( H_q(t, \epsilon) \) as the solution to (1.4), with \( H_q \) in place of \( H \) and \( K_q \) in place of \( K_0 \). By construction, the intertwining relation (1.5) holds with \( V_q \) and \( P_q \) in place of \( V \) and \( P \). The motivation for this construction is that it yields the estimate

\[
\|U_\epsilon(t, s) - V_q(t, s)\| \leq \int_s^t \|K_q(u, \epsilon) - K_{q-1}(u, \epsilon)\| du \leq |t - s| \beta_q \epsilon^{q+1},
\]

for small \( \epsilon \), where \( \beta_q \) is finite. When \( H(t) \) satisfies the analyticity hypothesis \( \mathbf{A} \), one can further control \( \beta_q \) as a function of \( q \) and prove the bound \( \beta_q \leq c_0 c_q^q q! \). The optimization over \( q \) can then be performed. This implies the following theorem.

**Theorem 2.4.** [38, 28] Assume \( H(t) \) satisfies \( \mathbf{A} \) and \( \mathbf{G} \). There exist \( \Gamma > 0 \) and an orthogonal projector \( P_\star(t, \epsilon) = P(t) + O(\epsilon) \) such that

\[
\|(I - P_\star(t, \epsilon))U_\epsilon(t, s)P_\star(s, \epsilon)\| = O(e^{-\Gamma/\epsilon}|t - s|).
\]

**Remarks 2.5.**

1. This result shows that although the transitions between instantaneous eigenprojectors of \( H(t) \) are of order \( \epsilon \) for finite times, there exists another set of orthogonal projectors which are better suited to the adiabatic approximation, in the sense that transitions between them are always exponentially small.
2. We call the projectors \( P_\star(t, \epsilon) \) optimal adiabatic projectors. They are not uniquely defined.
3. We recover Theorem 2.1 if we further assume \( \mathbf{S} \). See (2.3).
4. The construction above was used in [28] and [23] to reduce the study of transitions between two isolated levels in the spectrum of a general Hamiltonian to an effective reduced 2-level problem.

### 3. Time Development of Exponentially Small Non–Adiabatic Transitions

Berry [3] pushes these ideas further for two-level systems. For such systems, a one-dimensional projector corresponds to an orthonormal basis of \( \mathbb{C}^2 \), with one vector in the range and one vector in the kernel of the projector. Berry proposes studying the time development of the transition amplitude in this time-dependent basis. His heuristic arguments [3] are supported by spectacular numerical results [33]. These show that the transition amplitude passes from its initial value of zero to its final exponentially small value (2.1) in a universal way described by an error function.

The first rigorous results [20] on this question were obtained for the particular \( 2 \times 2 \) Hamiltonian function

\[
H(t) = \frac{1}{2\sqrt{\delta^2 + t^2}} \begin{pmatrix} \delta & t \\ t & -\delta \end{pmatrix},
\]

for fixed \( \delta > 0 \). The eigenvalues are \( E_1(t) = 1/2 \) and \( E_2(t) = -1/2 \) for every \( t \).

This Hamiltonian can be viewed as the Landau–Zener Hamiltonian modified to keep its eigenvalues constant and with the gap normalized to the value 1.

The notion of avoided crossing has been replaced by the singularities of the Hamiltonian at \( t = \pm i \delta \). These points govern the transitions between the two levels and yield the transition amplitude \( \sqrt{2} e^{-\delta/\epsilon} \). See [22]. The time development is described by the following result.
THEOREM 3.1. [20] Let \( H(t) \) be given by (3.1) and let \( \varphi_1(t) \) and \( \varphi_2(t) \) be smooth real-valued normalized eigenvectors corresponding to the eigenvalues 1/2 and \(-1/2\), respectively. Then for any \( \mu \in (0,1/2) \),

1. There exist vectors \( \chi_1(\epsilon, t) \) and \( \chi_2(\epsilon, t) \) that satisfy the Schrödinger equation (1.1) up to errors of order \( e^{-\delta/\epsilon} \) and correspond to the eigenstates \( \varphi_1(t) \) and \( \varphi_2(t) \) in the sense that

\[
\lim_{|t| \to \infty} |\langle \varphi_j(t), \chi_j(\epsilon, t) \rangle| = 1 + O(e^{-\delta/\epsilon} e^{\mu}).
\]

Moreover, the set \( \{ \chi_j(\epsilon, t) \}_{j=1,2} \) is orthonormal up to errors of order \( e^{-\delta/\epsilon} e^{\mu} \).

2. The Schrödinger equation has solutions \( \Psi_j(\epsilon, t), j = 1,2 \), such that uniformly in \( t \in \mathbb{R} \) as \( \epsilon \to 0 \),

\[
\Psi_1(\epsilon, t) = \chi_1(\epsilon, t) - \sqrt{2} e^{-\delta/\epsilon} \frac{1}{2} \left\{ \text{erf} \left( \frac{t}{\sqrt{2} \delta \epsilon} \right) + 1 \right\} \chi_2(\epsilon, t) + O(e^{-\delta/\epsilon} e^{\mu}),
\]

and

\[
\Psi_2(\epsilon, t) = \chi_2(\epsilon, t) + \sqrt{2} e^{-\delta/\epsilon} \frac{1}{2} \left\{ \text{erf} \left( \frac{t}{\sqrt{2} \delta \epsilon} \right) + 1 \right\} \chi_1(\epsilon, t) + O(e^{-\delta/\epsilon} e^{\mu}).
\]

REMARKS 3.2. 1. Recall that the function \( \text{erf} \) is defined by

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} \, dy \in [-1, 1].
\]

The vectors \( \chi_j(\epsilon, t), j = 1,2 \), are explicitly constructed as approximate solutions to (1.1) obtained by means of optimal truncation of asymptotic expansions of actual solutions. As \( t \to -\infty \) they are asymptotic to the instantaneous eigenvectors \( \varphi_j(t) \) of \( H(t) \), up to a phase. We call them the optimal adiabatic states.

2. The transition mechanism between optimal adiabatic states goes from the value zero to the value \( \sqrt{2} e^{-\delta/\epsilon} \) in a smooth monotonic way described by the switching function \( \text{erf}+1)/2 \), on a time scale of order \( \sqrt{T} \). By contrast, the transition between instantaneous eigenstates of the Hamiltonian displays oscillations whose amplitudes are of order \( \epsilon \) for any finite time. It reaches its exponentially small value only at \( t = \infty \). The vectors \( \chi_j(\epsilon, t) \) are also optimal in that sense.

3. Since the optimal adiabatic states and eigenstates essentially coincide at \( t = \pm \infty \), the transition amplitude equals \( A(\epsilon) \simeq \sqrt{2} e^{-\delta/\epsilon} \), up to errors of order \( e^{-\delta/\epsilon} e^{\mu} \).

4. The whole evolution operator \( U_\epsilon(t, s) \) associated with (1.2) is actually known up to errors of order \( e^{-\delta/\epsilon} e^{\mu} \), for any \( t \) and \( s \).

We prove this theorem by carefully controlling the terms that occur in the asymptotic expansion of exact solutions, so that we obtain very detailed information about the optimally truncated expansion. Even for our simple Hamiltonian, this involves very delicate asymptotic analyses of solutions to several non-linear recursion relations.

This result was soon followed by an essentially equivalent result by Betz and Teufel [5], who dealt with a slightly larger family of Hamiltonians. The main point of their analysis was to introduce a different technique. They defined the optimal adiabatic states as instantaneous eigenstates of transformed Hamiltonians, similar to (2.2), instead of dealing with the expansions of solutions. This way they simplified the analysis of the recursion relations. That paved the way for their
significant generalization [6]. They proved that the conclusions of Theorem 3.1 held for general $2 \times 2$ time dependent analytic Hamiltonians.

Let us informally state their result for real symmetric matrices

$$ H(s) = \frac{1}{2} \begin{pmatrix} Z(s) & X(s) \\ X(s) & -Z(s) \end{pmatrix}, $$

such that $\rho(s) = \sqrt{X^2(s) + Z^2(s)}$ is strictly positive for real $s$ and has only one zero $z_0 \in \mathbb{C}^+$, the upper half plane. Under certain technical assumptions, Betz and Teufel show that the conclusion to Theorem 3.1 holds with $t$ given by

$$ t(s) = \int_0^s \rho(u) \, du, $$

$\delta$ given by $\delta = |\text{Im} \, t(s)|_{s=z_0}$, and the prefactor $\sqrt{2}$ in front of the curly brackets replaced by $G \neq 0$, which depends on the properties of $H(s)$ in a neighborhood of $z_0$.

See [8] for a non-technical account of their findings, which confirm Berry’s intuitive arguments.

4. Exponentially Accurate Born–Oppenheimer Approximation

We now turn to transition phenomena in the Born-Oppenheimer approximation. This approximation is widely used in molecular dynamics. It takes advantage of the disparity between the masses of nuclei and electrons, and accordingly, of the very different time scales for these particles’ motions. The light electrons adjust quickly to the relatively slow changes in the configuration of the heavy nuclei. If they start in an electronic bound state, they stay in that bound state, as if the nuclei were not moving. This is typical adiabatic behavior. The nuclei are heavy, so they behave in a semiclassical fashion. The two motions are coupled because the electronic energy level plays the role of an effective potential for the nuclei.

Hence, the Born-Oppenheimer approximation is simultaneously adiabatic and semiclassical in nature. The breakdown of these approximations are both associated with exponentially small phenomena: non-adiabatic transitions on the one hand, and tunnelling on the other hand. See e.g., [16, 17, 11]. It is thus reasonable to expect the Born-Oppenheimer approximation to be valid up to exponentially small corrections.

We set the electron masses equal to one and take the nuclear masses proportional to $\epsilon^{-4}$, where $\epsilon$ is a small parameter. We let $x \in \mathbb{R}^d$ denote the collective positions of the nuclei and $y \in \mathbb{R}^n$ denote those of the electrons. If $W(x, y)$ is the inter-particle potential, the molecular Hamiltonian is

\begin{equation}
H(\epsilon) = -\frac{\epsilon^4}{2} \Delta_x - \frac{1}{2} \Delta_y + W(x, y),
\end{equation}

on the Hilbert space $L^2(\mathbb{R}^d \times \mathbb{R}^n)$, where $\Delta_z$ denotes the Laplacian in the $z$ variables. We introduce the electronic Hilbert space $\mathcal{H}_{el} = L^2(\mathbb{R}^n)$ and the electronic Hamiltonian $h(x) = -\frac{1}{2} \Delta_y + W(x, y)$ on $\mathcal{H}_{el}$, which depends parametrically on the position $x$ of the nuclei. Then we can rewrite (4.1) as

\begin{equation}
H(\epsilon) = -\frac{\epsilon^4}{2} \Delta_x + h(x)
\end{equation}
In an appropriate time scale, the molecular time-dependent Schrödinger equation is

\begin{equation}
\epsilon^2 \frac{\partial \Psi}{\partial t} = -\frac{\epsilon^4}{2} \Delta_x \Psi + h(x) \Psi, \quad \Psi \in L^2(\mathbb{R}^d, \mathcal{H}_{el}).
\end{equation}

We study solutions to this equation in the small $\epsilon$ limit, for times $t \in [0, T]$, where $T$ is fixed. We never actually use the explicit form of the Hamiltonian (4.1). We study solutions to (4.3) with (4.2) satisfying the following two hypotheses. The first one says $x \mapsto h(x)$ is analytic in an appropriate sense:

$A_{el}$: (i) For any $x \in \mathbb{R}^d$, $h(x)$ is a self-adjoint operator on some dense domain $\mathcal{D} \subset \mathcal{H}_{el}$, where $\mathcal{H}_{el}$ is the electronic Hilbert space. We assume the domain $\mathcal{D}$ is independent of $x$, and that $h(x)$ is bounded below, uniformly for $x \in \mathbb{R}^d$.

(ii) There exists $\mu > 0$, such that for every $\psi \in \mathcal{D}$, the vector $h(x)\psi$ is analytic in $S_\mu = \{z \in \mathbb{C}^d : \text{Im}(z) < \mu, \quad j = 1, \ldots, d \}$.

The second hypothesis is the familiar gap hypothesis, expressed this time for a space variable:

$G_{el}$: There exists an open set $\Xi \subset \mathbb{R}^d$, such that for all $x \in \Xi$, there exists an isolated, multiplicity one eigenvalue $E(x)$ of $h(x)$ associated with a normalized eigenvector $\Phi(x) \in \mathcal{H}_{el}$.

Under these two hypotheses, the Born–Oppenheimer approximation colloquially says the following: Pick an initial condition for (4.3) that is the product of a semiclassical wave packet concentrated near $(a(0), \eta(0))$ in phase space, times the electronic eigenstate $\Phi(x)$. Then, at some later time $t > 0$, the solution $\Psi$ is again given to leading order in $\epsilon$ by a semiclassical wave packet times the same electronic eigenstate. The nuclear wave packet is now concentrated near the phase space point $(a(t), \eta(t))$, determined from $(a(0), \eta(0))$ by the classical flow generated by the classical Hamiltonian $p^2/2 + E(x)$. The spreading of the nuclear wave packet is also determined by the classical dynamics.

As implied above, there is an exponentially accurate version of the Born-Oppenheimer approximation.

**Theorem 4.1.** [18] Assume hypotheses $A_{el}$ and $G_{el}$. There exists a vector $\Psi_\epsilon(x, t, \epsilon) \in L^2(\mathbb{R}^d, \mathcal{H}_{el})$ (that depends on a parameter $g$) for $t \in [0, T]$, such that for small values of $g$, there exist $C(g)$ and $\Gamma(g) > 0$, such that for small $\epsilon > 0$,

$$\left\| e^{-iH(\epsilon)/\epsilon^2} \Psi_\epsilon(x, 0, \epsilon) - \Psi_\epsilon(x, t, \epsilon) \right\|_{L^2(\mathbb{R}^d, \mathcal{H}_{el})} \leq C(g) e^{-\Gamma(g)/\epsilon^2}.$$

In the state $\Psi_\epsilon(x, t, \epsilon)$, the electrons have a probability of order $\epsilon^4$ to be in the electronic subspace $\text{Span} \{\Phi(x)\}^\perp$. For any $b > 0$ and sufficiently small values of $g$, the nuclei are localized near a classical path $a(t)$ in the sense that there exist $c(g)$ and $\gamma(g) > 0$, such that for small $\epsilon > 0$,

$$\left( \int_{|x-a(t)| > b} \left\| \Psi_\epsilon(x, t, \epsilon) \right\|^2_{\mathcal{H}_{el}} dx \right)^{1/2} \leq c(g) e^{-\gamma(g)/\epsilon^2}.$$
The nuclear configuration $a(t)$ is determined by classical dynamics in the effective potential $E(x)$. The limiting time $T > 0$ is determined by the condition $\{a(t), t \in [0,T]\} \subset \Xi$.

**Remarks 4.2.** 1. Under more restrictive hypotheses, the limiting time $T$ can be allowed to grow like $|\ln(\epsilon^2)|$, the Ehrenfest time scale. The error term becomes $e^{-\Gamma(\sigma)/\epsilon^\sigma}$, where $0 < \sigma < 2$. 2. The exponentially accurate approximation has non-trivial components in the electronic space orthogonal to the electronic eigenvector $\Phi(x)$ which are of order $\epsilon^2$, but its semiclassical dynamics is entirely determined by the electronic level $E(x)$. These components of order $\epsilon^2$ correspond to the adjustment of the spectral subspace necessary to achieve exponential accuracy at finite times. See (2.3).

We shall not be more precise here about the construction of the approximate solution $\Psi_*$. We do not need it, and the details can be found in [18]. We just mention that $\Psi_*$ is obtained by optimal truncation of an asymptotic series in powers of $\epsilon$ for the solution to (4.3). Results of the same sort have then been obtained in [39, 36, 42] in more general situations and in other contexts. Roughly speaking, these papers show that transitions between (adjusted) spectrally isolated electronic subspaces are exponentially small under the time evolution. They can handle nuclear wave packets which are not necessarily concentrated near a point of phase space in the semiclassical limit. In order to determine the characteristics of the nuclear component of the solutions, a semiclassical analysis within the spectral subspaces needs to be done. An important feature of these papers is that they decouple the adiabatic and the semiclassical effects.

In any case, we infer from these results that transitions between electronic subspaces are hidden in the exponentially small error term. They are expected to stem from some Landau–Zener mechanism. The results in the next section verify that this expectation is correct.

5. Non-Adiabatic Corrections to the Born–Oppenheimer Approximation

In Theorem 4.1, the nuclear wave packet is concentrated near $a(t)$. Thus, it is reasonable to expect the electrons to behave as though their motion were governed by $h(a(t))$. Furthermore, it is reasonable to expect the Landau–Zener mechanism to govern their transitions, since the replacement of $h(x)$ by $h(a(t))$ leads to a standard adiabatic approximation.

From the analysis [21] of a simple Born–Oppenheimer model that we discuss below, the main *qualitative* features of this argument are correct. However, the *quantitative* predictions based on the use of the Landau–Zener formula are wrong at leading order. For example, the exponential decay rate is wrong, even if the width of the nuclear wave packet is negligible in the limit $\epsilon \to 0$. Some characteristics of the energy/momentum density of the nuclear wave packet are present in the main features of the piece of the wave function that has made a transition. These aspects are clearly absent from the heuristic considerations above.

The following situation is considered in [21]: The nuclear variable $x$ belongs to $\mathbb{R}$ and the electronic Hamiltonian is given by a finite order matrix $h(x, \delta)$. We restrict attention to a $2 \times 2$ matrix for simplicity. We assume this matrix depends
analytically on $x$ and approaches limits sufficiently rapidly as $x \to \pm \infty$. As we saw previously, this is the most convenient framework in which a Landau–Zener formula can be proven. We finally assume that the two non-degenerate electronic levels display only one avoided crossing in a neighborhood of $(x, \delta) = (0, 0)$, of the type described by (1.8). Here again, $\delta$ is a supplementary parameter of the problem. More precisely, with the notation of (5.3), we assume:

**MC:** i) For some $\delta_0 > 0$ and $\mu > 0$, the map $S_\mu \times [0, \delta_0] \ni (x, \delta) \mapsto h(x, \delta) \in M_2(\mathbb{C})$ is $C^2$. For any fixed $\delta$, $x \mapsto h(x, \delta)$ is analytic in $S_\mu$.

ii) There exist $\nu > 5/2$, $c \geq 0$, and two matrices $h(\pm \infty, \delta)$, such that

$$\sup_{|y| \leq \mu, \delta \in [0, \delta_0]} \| h(x + iy, \delta) - h(\pm \infty, \delta) \| \leq \frac{c}{(x) \nu}, \quad \text{as } x \to \pm \infty.$$  

The limiting matrices $h(\pm \infty, \delta)$ are $C^2$ in $\delta \in [0, \delta_0]$ and are non-degenerate.

For each $x \in \mathbb{R}$ and each $\delta \in [0, \delta_0]$, we denote the two eigenvalues of $h(x, \delta)$ by $E_1(x, \delta)$ and $E_2(x, \delta)$. We assume:

**AC:** i) For $\delta > 0$ and $x \in [-\infty, +\infty]$, the eigenvalues satisfy $E_1(x, \delta) < E_2(x, \delta)$.

ii) When $\delta = 0$, the analytic eigenvalues $E_j(x, 0)$ have only one real crossing at $x = 0$. They are labeled so that $E_1(x, 0) < E_2(x, 0)$ for all $x < 0$.

iii) Moreover, we assume the following behavior near the avoided crossing:

$$E_2(x, \delta) - E_1(x, \delta) = \sqrt{x^2 + \delta^2 + R_3(x, \delta)},$$

where $R_3$ is a remainder of order 3 around $(0, 0)$.

The corresponding normalized eigenvectors are denoted by $\phi_1(x, \delta)$ and $\phi_2(x, \delta)$. They are chosen to satisfy the phase condition $\langle \phi_j(x, \delta), \partial_x \phi_j(x, \delta) \rangle \equiv 0$.

**Remark 5.1.** In the following discussion, $\delta$ is positive and fixed, so we often drop it from the notation. We reintroduce it when necessary.

The analysis of (4.3) in this set-up starts with a separation of variables and the study of generalized eigenvectors $\Phi$ of $H(\epsilon)$. These vectors depend on a real energy parameter $E$, and they are defined by the ODE in $\mathbb{C}^2$:

$$\left( -\frac{\epsilon^4}{2} \partial_x^2 + h(x) \right) \Phi = E \Phi.$$

We pick $E$ in some bounded energy interval $\Delta \subset \mathbb{R}$. To ensure that scattering is possible, we assume

$$(5.2) \quad \inf_{E \in \Delta} \Delta > \sup_{x \in \mathbb{R}} E_2(x).$$

Choosing an energy density $Q(E, \epsilon) \in \mathbb{C}$, we get an exact solution to (4.3) of the form

$$(5.3) \quad \int_{\Delta} Q(E, \epsilon) \Phi(x, E, \epsilon) e^{-itE/\epsilon^2} dE.$$

From physical intuition, we anticipate that if the energy density is sharply peaked, then when $|t|$ is large, these solutions to (4.3) will be concentrated in a region where $|x|$ is large.
We choose the energy density to behave like a Gaussian of width $\epsilon^2$, centered around an initial energy $E_0$ in the interior of $\Delta$. More precisely we assume:

**ED:** The complex valued energy density is supported on $\Delta$ and has the form

$$Q(E, \epsilon) = e^{-G(E)/\epsilon^2} e^{-i J(E)/\epsilon^2} P(E, \epsilon),$$

where

i) $G \in C^3(\Delta)$ is non-negative, and satisfies $G(E) = g(E-E_0)^2/2 + O(E-E_0)^3$, for some $E_0$ in the interior of $\Delta$ and some $g > 0$.

ii) $J \in C^3(\Delta)$ is real-valued.

iii) $P(E, \epsilon) \in C^1(\Delta)$ satisfies $\sup_{E \in \Delta, \epsilon \geq 0} |\partial_E^n P(E, \epsilon)| \leq C_n$, for $n = 0, 1$.

This choice implies that in the remote past or remote future the nuclear components are coherent states which basically propagate freely along the different electronic levels. Modulo the technicality of inserting a cut-off, Gaussian coherent states have such an energy density. For these states we use the notation of [15]:

$$\varphi_0(A, B, \epsilon^2, a, \eta, x) = \frac{1}{\pi^{1/4} \epsilon^{1/2} A^{1/2}} \exp \left( -\frac{B(x-a)^2}{2 A \epsilon^2} + i \frac{\eta (x-a)}{\epsilon^2} \right),$$

where the complex numbers $A$ and $B$ satisfy the normalization condition $\text{Re} B A = 1$. These states are localized in position near $x = a$, and in momentum near $p = \eta$. Their position uncertainty is $\epsilon |A|$ and their momentum uncertainty is $\epsilon |B|$. For a thorough discussion of these wave packets, see [15]. Note, however, that there are states characterized by a Gaussian energy density which are not necessarily Gaussian in space and momentum. See Section 6 of [21].

The solutions of (4.3) considered in [21] are characterized by their behavior in the remote past. We now describe these asymptotic states. The generalized eigenvector equation (5.1) can be solved by means of a WKB Ansatz as follows: Let $k_j(x, E) = \sqrt{2(E - E_j(x))}$, $j = 1, 2$, be the classical momenta corresponding to the potentials $E_j(x)$. Then, the solutions $\Phi(x, E, \epsilon)$ to the stationary equation (5.1) can be written as

$$\Phi(x, E, \epsilon) = \sum_{j=1,2, \sigma = \pm} \phi_j(x) \frac{c_j^\sigma(x, E, \epsilon)}{\sqrt{2 k_j(x, E)}} e^{-i \sigma \int_0^x k_j(y, E) dy/\epsilon^2},$$

where $c_j^\sigma(x, E, \epsilon) \in \mathbb{C}$ are coefficients that satisfy some linear ODE in $x$, for each $E \in \Delta$. The condition $\mathcal{M}_L \Phi$ ensures the existence of the limits of $\phi_j(x), k_j(x, E)$, and $c_j^\sigma(x, E, \epsilon)$, as $x \to \pm \infty$ and further provides rates at which they are approached.

We assume without loss that the “initial” conditions for the coefficients $c_j^\sigma$ are fixed at $x = -\infty$ and that $c_j^\sigma(-\infty, E, \epsilon)$ are independent of $\epsilon$ and $E$. It follows that the dependence in $E$ of all the above quantities is analytic in a neighborhood of $\Delta$, and $c_j^\sigma(x, E, \epsilon)$ is uniformly bounded in $(x, E, \epsilon)$.

We now introduce freely propagating states $\psi^\sigma(x, t, \epsilon, \pm) \in L^2(\mathbb{R}, \mathbb{C}^2)$ that describe the asymptotics of the solutions $\Psi(x, t, \epsilon)$ of (4.3) as $t \to \pm \infty$. Let

$$\psi^\sigma(x, t, \epsilon, \pm) = \sum_{j=1, 2} \phi_j(x) \int_\Delta \frac{Q(E, \epsilon) c_j^\sigma(\pm \infty, E, \epsilon)}{\sqrt{2 k_j(\pm \infty, E)}} e^{-itE/\epsilon^2} e^{-i \sigma (zk_j(\pm \infty, E) + \omega_j^2(E))/\epsilon^2} dE,$$
where $\omega_j^\pm(E) = \lim_{x \to \pm\infty} \int_0^x k_j(y, E) \, dy - x k_j(\pm\infty, E)$. These states are linear combinations of products of free scalar wave packets in constant scalar potentials times eigenvectors of the electronic Hamiltonian. Their propagation is thus governed by the various channel Hamiltonians. With our sign convention, the states indexed by $\sigma = -$ propagate to the right, whereas those indexed by $\sigma = +$ travel to the left.

By construction, $\psi^+(x, t, \epsilon, \pm)$ is likely to be a good approximation of $\Psi(x, t, \epsilon)$ only for $x \to \pm\infty$, i.e., for $t \to \pm\infty$. Similarly, $\psi^+(x, t, \epsilon, \pm)$ should be a good approximation only for $x \to \pm\infty$, i.e. for $t \to \mp\infty$. Quantitatively, we have

**Proposition 5.2.** Assume $\mathcal{AC}_{el}$, $\mathcal{M}_{el}$ and $\mathcal{ED}$. As $t \to \pm\infty$,

$$\| \Psi(\cdot, t, \epsilon) - \psi^\sigma(\cdot, t, \epsilon, \mp \sigma) \|_{L^2(\mathbb{R}^+)} = O(1/|t|^\beta),$$

for any $0 < \beta < 1/2$.

**Remark 5.3.** The estimate above is valid for any value of $\epsilon > 0$, and its proof essentially relies on integration by parts.

We now analyze the freely propagating states $\psi^\sigma(x, t, \epsilon, \pm)$. We do this by investigating the small $\epsilon$ behavior of the coefficients $c_j^\sigma(\pm\infty, E, \epsilon)$. We are interested in solutions associated with only one electronic state in the remote past. For concreteness, we choose level 2. Proposition 5.2 shows that this corresponds to the initial condition

$$c_j^\sigma(-\infty, E, \epsilon) = \delta_{j,2} \delta_{\sigma,-}.$$  

(5.7)

The component that has made the transition to level 1 in the course of the evolution, and propagates in the positive direction for large positive times is $\psi^+_1(x, t, \epsilon, +)$. It is characterized by $c_j^+(+\infty, E, \epsilon)$. We analyze solutions of the equation satisfied by $c_j^\sigma$ by means of complex WKB methods, along the lines of [24], since we are in an avoided crossing regime. At this step, we require the parameter $\delta > 0$ to be small, so that we can apply the results of [24]. They show that there exist $\delta_0 > 0$, and $\delta \to \delta_0(\delta)$, such that for $\delta < \delta_0$ and $\epsilon < \epsilon_0(\delta)$,

$$c_j^+(+\infty, E, \epsilon) = e^{-i\theta(\zeta)} e^{\int_{C_0} k_j(E, z) \, dz/e^2} (1 + O(\epsilon^2)),$$

(5.8)  

for all $E \in \Delta$. Here $\zeta$ is a loop in the complex plane, based at the origin, which encircles the zero $z_0(\delta) \in \mathbb{C} \setminus \mathbb{R}$ of $k_1(z, E) - k_2(z, E)$ closest to the real axis. Also, $\theta(\zeta) \in \mathbb{C}$ only depends on the analytic continuation along $\zeta$ of the electronic eigenvectors.

We now have everything to state the main result of [21]:

**Theorem 5.4.** Let $\psi(x, t, \epsilon)$ be a solution to the molecular Schrödinger equation (4.3) with electronic Hamiltonian $h(x, \delta)$ satisfying hypotheses $\mathcal{M}_{el}$ and $\mathcal{AC}_{el}$. Assume the solution is characterized asymptotically in the past for negative $x$'s by

$$\lim_{t \to -\infty} \| \psi(x, t, \epsilon) - \psi^\sigma_2(x, t, \epsilon, -) \|_{L^2(\mathbb{R}^-)} = 0,$$

with

$$\psi^\sigma_2(x, t, \epsilon, -) = \phi_j(x) \int_\Delta \frac{Q(E, \epsilon)}{\sqrt{2k_2(-\infty, E)}} \, e^{-it E/e^2} e^{i(xk_2(-\infty, E) + \omega_j^+(E))/e^2} \, dE,$$

where the energy density $Q(E, \epsilon)$ satisfies $\mathcal{ED}$. Let $E(k) = k^2/2 + E_1(+\infty)$

$$\alpha(E) = G(E) + \text{Im} \int_\zeta k_2(z, E) \, dz,$$

$$\kappa(E) = J(E) - \text{Re} \int_\zeta k_2(z, E) \, dz - \omega_j^+(E).$$
Assume $E^*$ is the unique absolute minimum of $\alpha(\cdot)$ and set $k^* = \sqrt{2(E^*-E_1(\infty))}$.

Then, there exist $\delta_0 > 0$, $p > 0$ arbitrarily close to $5/2$, and a function $\epsilon_0 : (0, \delta_0) \to \mathbb{R}^+$, such that for all $0 < \beta < 1/2$, $\delta < \delta_0$, and $\epsilon < \epsilon_0(\delta)$, the following asymptotics hold as $t \to \infty$:

$$\psi_1(x,t,\epsilon,+) = \frac{\chi(x) e^{-i\theta(\epsilon)/2}{\pi}^{3/4} e^{-\alpha(E^*)/t^2}}{(\frac{\pi}{2\alpha(E(k))})^{1/4}} e^{iS_+(t)/t^2} \varphi_0(A_+(t), B_+, \epsilon^2, a_+(t), \eta_+, x)$$

$$\times P(E^*,\epsilon) \sqrt{k^*} e^{-i(\alpha(E^*)-k^*^2\kappa(E^*))/t^2} + O(e^{-\alpha(E^*)/t^2}) + O_{\epsilon}(1/t^3),$$

where $\varphi_0$ is a Gaussian (5.4) parametrized by

$$\eta_+ = k^*, \quad a_+(t) = \kappa(E^*) + k^* t, \quad B_+ = 1/\sqrt{\partial^{2}_{k\kappa}(E(k))|_{k=k^*}}, \quad A_+(t) = B_+ \left( \partial_{k}^{2}\alpha(E(k))|_{k=k^*} + i(\partial_{k\kappa}^{2}\kappa(E(k))|_{k=k^*} + t) \right), \quad \text{and}$$

$$S_+(t) = \left( k^*^2/2 - E_1(\infty) \right) t.$$ 

All error terms are estimated in the $L^2(\mathbb{R})$ norm, and the estimate $O(e^{-\alpha(E^*)/t^2})$ is uniform in $t$. The $O_{\epsilon}(1/t^3)$ may depend on $\epsilon$.

Remark 5.5. The leading order of the piece transmitted to level 1 is always given by an exponentially small prefactor times a freely propagating Gaussian nuclear wave packet, i.e., a Gaussian semiclassical wave packet in the constant potential given by the asymptotic electronic level. This is true even if the incoming state is not Gaussian.

All quantities computed from the electronic Hamiltonian depend on $\delta$, despite this not being emphasized in the notation. Although we shall not do it here, one can further investigate the characteristics of the asymptotic state for small $\delta$'s.

Without being too precise, we mention that if $k_\epsilon(E) = \sqrt{2(E-E_2(0,0))}$ is the classical momentum at the crossing point of the electronic levels for $\delta = 0$, one has

$$\alpha(E) = g(E-E_0)^3/2 + \frac{\delta^2\pi}{4k_\epsilon(E)} + O((E-E_0)^{3}) + O(\delta^3).$$

The factor $\delta^2\pi/(4k_\epsilon(E))$ is the Landau-Zener decay rate computed from the electronic levels divided by the classical momentum at the avoided-crossing, and $g > 0$ stems from the energy density. So, for small values of $\delta$ and for a sufficiently narrow energy window $\Delta$, $\alpha$ is a quadratic term plus a positive, decreasing, convex function of $E$. This implies that $E^*$, determined by $\alpha'(E^*) = 0$, is strictly larger than $E_0$ and depends explicitly on the energy density. Furthermore, the decay rate has $\alpha(E^*) < \alpha(E_0)$. This has the following consequences.

Remark 5.6. The average momentum $k^*$ is larger than what a naïve application of energy conservation yields, and the exponentially small prefactor is larger than the prediction given by the Landau-Zener formula for $h(a(t))$. The correct values of both of these quantities depend explicitly on the chosen energy density.

The lesson we learn from this is that the leading behavior of the piece of the wave function that describes transitions from one electronic level to the other cannot be determined by straightforward semiclassical considerations and the Landau-Zener formula. The way the incoming asymptotic state is prepared in the remote past plays a crucial role in the main characteristics of the asymptotics.
Results of this sort are obtained in [26] for scattering systems defined by more general semiclassical autonomous PDEs entertaining dispersive waves. While the overall general strategy remains the same, the separation of variables, study of the stationary equation by means of complex WKB methods and analysis of the space-time properties of exact solutions, can be quite different.

The description of the time development of the transitions between the electronic levels represents a challenging question. Betz and Teufel [7] have heuristic arguments and numerical evidence supporting a description of this phenomenon. However, the mathematical problem remains open.

References

Department of Mathematics and Center for Statistical Mechanics and Mathematical Physics, Virginia Tech, Blacksburg, Virginia, 24060–0123, USA
E-mail address: hagedorn@math.vt.edu

Institut Fourier, Unité Mixte de Recherche CNRS–UJF 5582, Université de Grenoble I, BP 74, F–38402 Saint Martin d’Hères Cedex, France
E-mail address: alain.joye@ujf-grenoble.fr